The MODIS 500 Meter Global Vegetation Continuous Field Products

Matthew C. Hansen, Ruth S. DeFries, John R. G. Townshend, Mark Carroll, Charlene Dimiceli and Rob A. Sohlberg

University of Maryland, Department of Geography, 2181 LeFrak Hall, College Park, MD 20742
Tel: 301-314-2585, Fax: 301-314-9299, Email: mhansen@geog.umd.edu

1. Introduction

Global maps are useful for a variety of applications, including biogeochemical modeling exercises as well as deriving natural resource inventories [1,2]. The state of knowledge of the earth surface has been improving over the last decade due to the increased use of synoptic satellite data sets to map global land cover. Most approaches to mapping large areas utilize an unsupervised classification approach [3,4]. Some researchers believe that unsupervised approaches are more appropriate as size of study areas increases to regional scales and beyond [5]. This is due to the difficulty in obtaining a robust training data set at these scales to drive spectral characterizations. However, others have attempted a supervised classification method using flexible algorithms such as decision trees which can easily delineate complex, multi-modal spectral signatures which are present within large study areas [6,7,8]. A major goal of global land cover mapping is having the ability to monitor land cover intermittently in order to measure land cover change. This requires a repeatable method which operates at the global scale. Of the two major classification approaches, unsupervised methods are the least amenable to automation as an analyst is required at each map iteration to label output spectral clusters. Supervised techniques are more easily automated than unsupervised techniques as once a training data set is derived, it can be used repeatedly to produce successive maps. Thus, for long term change monitoring, a supervised approach has distinct advantages.

However, global-scale classifications are inappropriate for change monitoring purposes. Measuring change using consecutive classifications has been shown to be of limited value in change detection studies. This problem is accentuated at the global scale. Global land cover characterizations are almost exclusively the domain of moderate and coarse spatial resolution sensors. As such, most pixels represent mixtures of constituent land cover categories. Given this reality, measuring change or repeating inventories using successive land cover classifications is problematic, unless some additional layer expressing probability of class membership is employed. Instead of using discrete categorical depictions of the land surface, depictions of coarse pixels as mixtures of constituent components allows for repeated inventories with the prospect of measuring change. Thus, for global monitoring an operational approach to repeated inventorying of land cover using sub-pixel estimates is the best choice.

This paper presents the first results of the Vegetation Continuous Fields (VCF) algorithm from the MODIS (Moderate Resolution Imaging Spectroradiometer) land cover suite of products. The methodology to derive these products fits the generic description of an operational global monitoring algorithm. It is supervised and derives proportional cover estimates suitable for change analyses. The mapping methodology is based on the prior efforts of DeFries et al. [9] and Hansen et al. [10]. Unlike other attempts to map the earth surface, it operates on the entire globe at once. Most other global products are actually individual continental-scale maps which have been stitched together. Characterizing the globe as a whole enhances the internal consistency of the final map.

The VCF products are global map layers representing proportional estimates of basic vegetation properties. The basic layers include percent tree cover, percent herbaceous/shrub cover and percent bare ground. For tree cover, percent leaf type (needleleaf and broadleaf) and percent leaf longevity (percent evergreen and percent deciduous) are also included. The continuous field methodology preserves detail by improving the depiction of heterogeneous landscapes as compared to classifications. It also allows for direct comparison between subsequent products at later dates and, therefore, the measurement of land cover change. This method for change detection has been prototyped for detecting long-term forest change globally based on AVHRR data [11,12]. Additional thematic layers are being developed to enhance the maps' usefulness, such as disturbed/undisturbed forest layers for the tropics. The initial products reveal a dataset of unprecedented detail for a global characterization and should be of use to a wide array of users, including biogeochemical modeling groups, regional land managers, and conservationists.

2. Data

The MODIS inputs consist of 8-day minimum blue reflectance composites. For this initial effort, the data were converted to 40 day composites using a second darkest albedo (sum of blue, green and red bands) algorithm to reduce the presence of cloud shadows. The inputs date from October 31, 2000 to December 9, 2001. An extra forty day composite period was added due to data gaps resulting from temporary sensor outages. The approach to mapping follows that of Hansen et al. [10] which employs annual phenological metrics derived from the MODIS composites to estimate percent tree cover. Multi-temporal metrics are appropriate for mapping global land cover as they are independent of the specific timing of vegetation dynamics. They are designed to work generically at the continental and global scale by capturing the salient points in the vegetation phenology which are shared by similar land cover types without regard to specific time of year. Metrics used in the VCF algorithm are listed in Table 1. The MODIS metrics are trained using reference data derived from high-resolution classifications aggregated to MODIS 500 meter cells. This yields the continuous percent cover training data.

Table 1. List of annual phenological metrics used in the VCF algorithm.

Metrics (all derived from 40-day composites) –
For bands 1-7 –
Minimum reflectance
Eighth darkest reflectance
Amplitude of minimum and 5th darkest reflectance
Mean 3 darkest reflectances
Mean 5 darkest reflectances
Mean 8 darkest reflectance
Reflectance at peak NDVI
Mean reflectance of values corresponding to 3 greenest composites
For NDVI –
Maximum NDVI
Eighth highest NDVI
Amplitude of minimum and 5th highest NDVI
Mean of 3 highest NDVI values
Mean of 5 highest NDVI values
Mean of 8 highest NDVI values
62 total inputs

3. Methods

The training data and metrics are input to a regression tree algorithm, with subsequent stepwise regression and bias adjustment steps per Hansen et al. [10]. The regression tree takes the following form:

$$D = \sum_{cases(j)} (y_j - u_{[j]})^2$$

where D is the deviance as measured by the corrected sum of squares for a split. This is calculated from all j cases of y and the mean value of those cases, u. All input satellite data are analyzed across digital number values and right and left splits are examined. The split which produces the greatest reduction in the residual sum of squares, or deviance, is used to divide the data and the process begins again for the two newly created subsets.

Figure 1 shows an example of data inputs and product for the percent tree cover map. Similar approaches are used for the other layers in a hierarchical fashion. The regression tree was used with the metrics to predict tree cover based on a training data set of 271 749 pixels. Regression trees are very robust and can fit to noise. In order to better generalize the relationships between dependent and independent variables, a two-step procedure is undertaken. First, a regression tree is grown using a 50 percent sample of the training data. The remaining half of training pixels are then used to prune, or better generalize the initial tree. These set aside pixels are run down the tree structure and the initial tree is pruned based on where additional nodes reduce the overall pruning data sum of squares by less than .01 percent. This final tree is then run on the global data and the mean values of the training data present in each node are assigned to the output 500 meter pixels. A stepwise regression is then run on each node and a bias adjustment is also performed for nodes exhibiting skewed distributions. These two final steps represent refinements to the initial regression tree node estimates and provide for a more continuous final map.

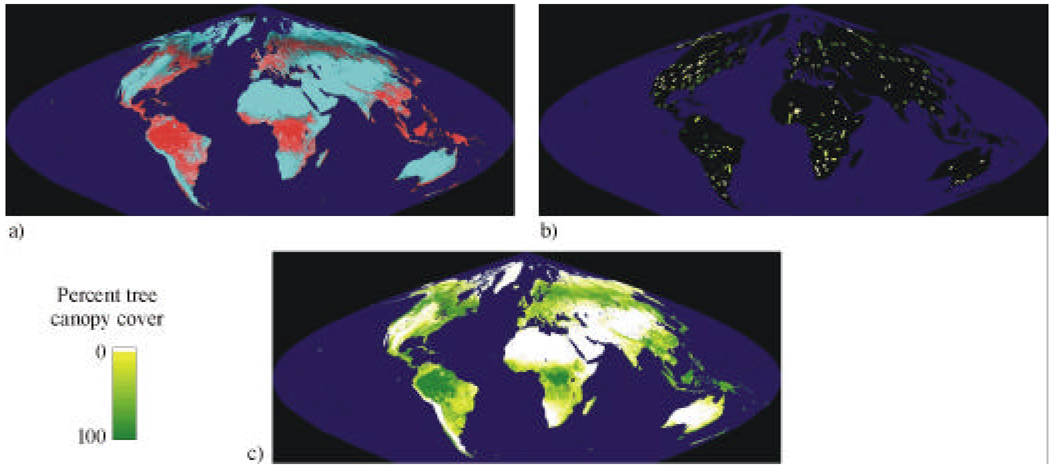
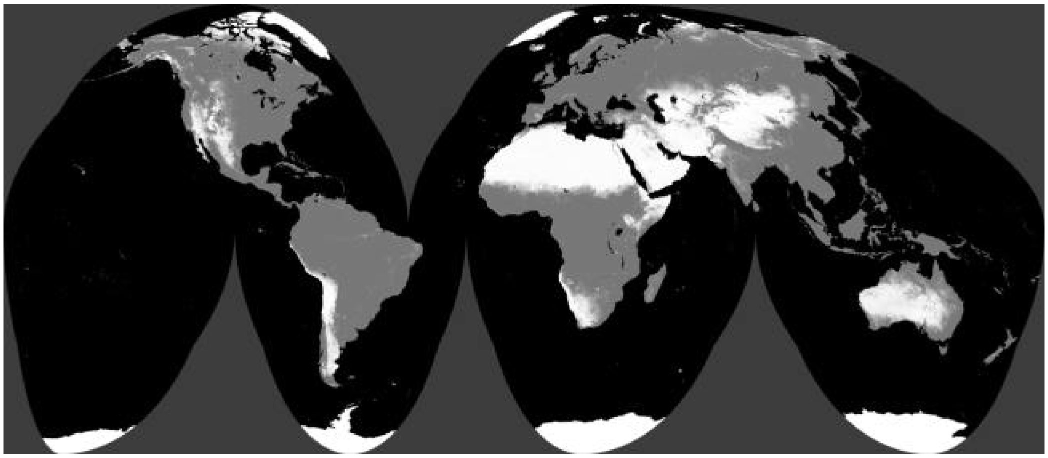


Figure 1. a) Example of two phenological metrics, where red is the mean NDVI of the 9 greenest months, and cyan is the mean red reflectance of the greenest 6 months. This first MODIS product employed 68 such metrics as independent variables in predicting tree cover. b) Training sites where each 500 meter pixel is labeled a percent tree cover value based on interpretations of high-resolution Landsat imagery. c) The final product derived using a regression tree algorithm.

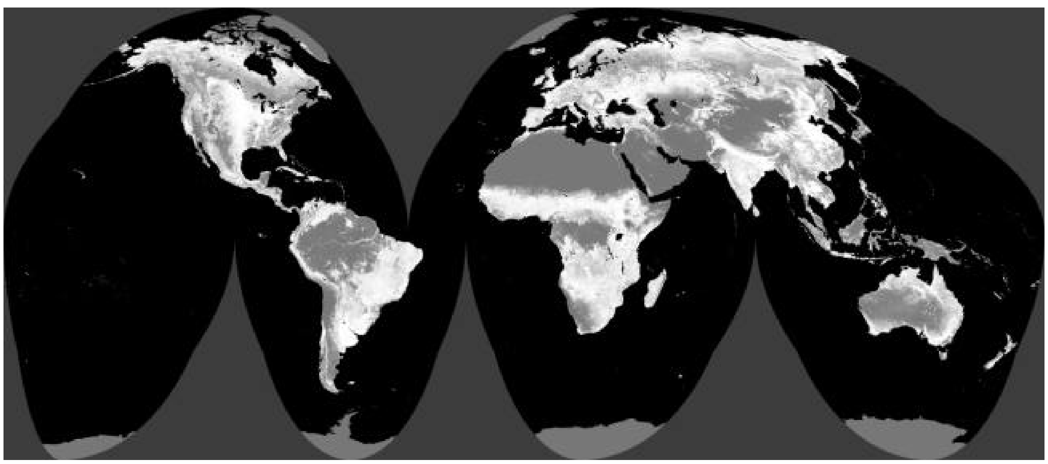
4. Results

Figure 2 shows the percent tree cover layer along with percent herbaceous/shrub and percent bare ground layers. Initial impressions of these data show an increased sensitivity in MODIS data for mapping low vegetation conditions. The northern limit of the Sahel is mapped without geologic features in the central Sahara appearing vegetated, as often occurred in AVHRR depictions. Additional layers to separate the components of the short vegetation types are being developed.

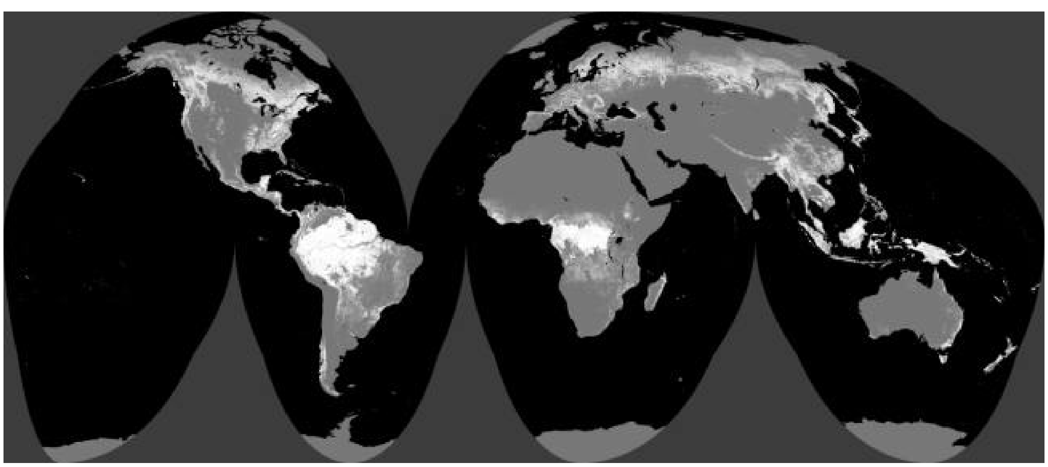
Initial results also reveal a tree cover layer which surpasses AVHRR-derived products in terms of spatial detail. Figure 3 shows an example of the MODIS tree cover map compared to an AVHRR depiction for an area in South America. The apparent blur in the AVHRR signal is due to many factors including broader bandwidths, Normalized Difference Vegetation Index (NDVI) compositing artifacts, among others. These factors tend to overemphasize the forest in the depiction. The MODIS data are more detailed, the result of the engineering of the bands for land monitoring as well as the fact that the 500 meter red and near infrared bands are derived from 250 meter pixels, greatly reducing adjacency effects in the 500 meter signal. Qualitative indications for all layers indicate greater spatial detail and improved spectral differentiation of cover categories.



a)



b)



c)

Percent cover

Figure 2. a) Percent bare ground, b) percent short vegetation and c) percent tall vegetation layers.

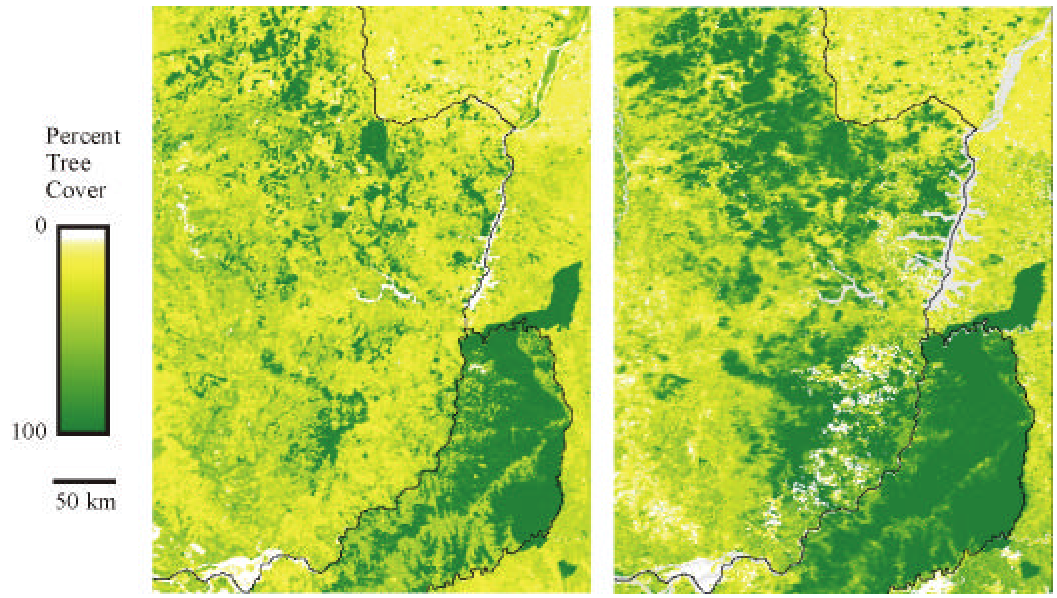


Figure 3. A subset of the MODIS VCF of percent tree cover for an area in South America centered at 55.38 degrees West, 25.35 degrees South. Country borders are in black. Paraguay takes up the majority of the image in the west, Brazil is east and Argentina south.

5. Validation

Each continuous field layer is tied to a single physiognomic vegetation trait. This allows for the implementation of a straightforward field validation protocol. For example, the VCF percent tree cover layer defines a tree as any woody plant greater than or equal to 5 meters in height. Canopy cover is defined as the percent of sky light in a plane orthogonal to the ground which is intercepted by plants defined as trees. Thus, a measure of the binomial distribution of canopy cover can be obtained in the field and directly compared to the satellite-derived estimate. Current work is ongoing in validating the percent tree cover layer. Field visits in conjunction with very-high resolution IKONOS imagery and high resolution Enhanced Thematic Mapper Plus data are used to create validation layers of percent cover for large test areas. Two such areas, one in Western Province, Zambia and another in the state of Colorado in the United States have been completed. Results are shown in Table 2. Each shows reasonable agreement with the VCF product. The method for creating the validation test area data sets is described in Hansen et al. [13]. Creating a number of sites across the globe as test areas for validating continuous field maps is a goal of future work.

Table 2. Standard error of estimate values at 500 meter spatial resolution for global VCF percent tree cover map compared with two validation test areas. Each test area was derived using field data and multi-resolution satellite imagery.

Validation Percent Tree Canopy Cover Strata	**<10**	**11-40**	**40-60**	**>60**	**Overall**
Colorado, USA WRS 035/032	5.3	13.3	13.0	11.7	11.6
Western Province, Zambia WRS 175/070-071	10.9	12.9	14.7	12.1	11.5

6. Discussion and conclusion

The 500 meter MODIS VCF layers represent a marked advance in the global mapping of land cover. The moderate resolution of the MODIS data fills an intermediate step between high and coarse resolution data and the spatial information gained is significant. Visible in the 500 meter map are individual clearings in the Amazon, fire history in the boreal zone, relic forests in montane regions of Africa, among other features not readily captured at coarser resolutions. The 250 meter MODIS data are currently being processed for the entire globe. Future maps will employ these data, providing more spatial detail and affording the opportunity to map at a scale more appropriate for viewing land cover change [14]. Validation work is evolving as well, particularly for the percent tree cover layer. Sites representative of global biomes are being chosen and studied with multi-resolution satellite data and field data to create coarse resolution validation test areas.

Acknowledgement

This research was supported by the National Aeronautics and Space Administration under contract NAS596060 and grant NAG59339.

References

[1] DeFries, R. S., Houghton, R. A., Hansen, M. C., Field, C. B., Skole, D. and Townshend, J., 2002, Carbon emissions from tropical deforestation and regrowth based on satellite observations for the 1980's and 1990's, *Proceedings of the National Academy of Sciences*, 99, 14256-14261.

[2] Matthews, E., 2001, Understanding the FRA, World Resources Institute Forest Briefing No. 1, (World Resources Institute: Washington, D. C.).

[3] Loveland, T. R., 2000, Development of a global landcover characteristics database and IGBP DISCover from 1km AVHRR data, *International Journal of RemoteSensing*, 21, 1303-1330.

[4] Belward A.S., E. Bartholomé, F. Achard, C. Carmona-Moreno, H. Eva, J-M. Grégoire, P. Mayaux and H-J. Stibig, 2002, Mapping Global Land Cover for the year 2000, *International Journal of Remote Sensing*, submitted.

[5] Cihlar, J., 2000, Land cover mapping of large areas from satellites: status and research priorities, *International Journal of RemoteSensing*, 21, 1093-1114.

[6] DeFries, R. S., Hansen, M. C., Townshend, J. R. G. and Sohlberg, R. S., 1998, Global land cover classifications at 8km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers, *International Journal of RemoteSensing*, 19, 3141-3168.

[7] Hansen, M. C., DeFries, R.S., Townshend, J.R.G., and Sohlberg, R., 2000, Global land cover classification at 1km spatial resolution using a classification tree approach *International Journal of Remote Sensing* 21, 1331-1364.

[8] Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., and Schaaf, C., 2002, Global land cover mapping from MODIS: Algorithms and early results, *Remote Sensing of Environment*, 83, 287-302.

[9] DeFries, R. S., Hansen, M. C., Townshend, J. R. G., Janetos, A. C., and Loveland, T. R., 2000, A new global 1-km dataset of percentage tree cover derived from remote sensing. *Global Change Biology*, 6, 247-254.

[10] Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Sohlberg, R. A., Dimiceli, C., and Carroll, M., 2002, Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data, *Remote Sensing of Environment*, 83, 303-319.

[11] Hansen, M. C., Chan, J. C.-W., Pagis, J., DeFries, R., and Luo, D., Long term change detection using continuous fields of tree cover from 8km AVHRR data for the years 1982-2000, in Analysis of Multi-Temporal Remote Sensing Images, vol. 2, eds. L. Bruzzone and P. Smits, (World Scientific: New Jersey), 363-370.

[12] Hansen, M. C., and DeFries, R. S., 2003, Using continuous fields of tree cover to map change from 8km AVHRR data for the years 1982-2000, submitted, *Ecosystems*.

[13] Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Marufu, L., Sohlberg, R., 2002, Development of a MODIS tree cover validation data set for Western Province, Zambia, Remote Sensing of Environment, 83, 320-335.

[14] Townshend, J. R. G. and Justice, C. O., 1988, Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations, *International Journal of Remote Sensing*, 9, 187-236.